AFRL-AFOSR-VA-TR-2015-0278

DERIVATIVE FREE OPTIMIZATION OF COMPLEX SYSTEMS **WITH THE USE OF STATISTICAL MACHINE LEARNING MODELS**

Katya Scheinberg
LEHIGH UNIVERSITY

04/06/2015
Final Report

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Executive Services, Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)* 12-09-2015	2. REPORT TYPE Final Performance	3. DATES COVERED *(From - To)* 15-08-2011 to 14-08-2014

4. TITLE AND SUBTITLE DERIVATIVE FREE OPTIMIZATION OF COMPLEX SYSTEMS WITH THE USE OF STATISTICAL MACHINE LEARNING MODELS	5a. CONTRACT NUMBER
	5b. GRANT NUMBER FA9550-11-1-0239
	5c. PROGRAM ELEMENT NUMBER 61102F
6. AUTHOR(S) Katya Scheinberg	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) LEHIGH UNIVERSITY 526 BRODHEAD AVE BETHLEHEM, PA 180153008 US	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AF Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/AFOSR RTA2
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
A DISTRIBUTION UNLIMITED: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This project focused on development of novel derivative free optimization methods that rely on recent techniques and models from statistical learning. The main idea of these methods is to build local models of the objective function from randomly sampled data points. This approach has many benefits, in that it allows us to construct fairly accurate models with relatively small number of samples. The key difference with the deterministic sampling approaches is that these accurate models are constructed with some high probability, but not always. Moreover, it is not known, when these models are accurate. Only the probability of an accurate model occurring is known. Under these conditions, novel convergence theory needed to be developed, which has been the focus of our research.

15. SUBJECT TERMS
optimization, Derivative-Free Optimization, Statistical Machine Learning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Katya Scheinberg
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU		
Unclassified	Unclassified	Unclassified			19b. TELEPHONE NUMBER *(Include area code)* 610-758-4039

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contract/Grant Title: DERIVATIVE FREE OPTIMIZATION OF COMPLEX SYSTEMS WITH THE USE OF STATISTICAL MACHINE LEARNING MODELS Contract/Grant #: FA9550-11-1-0239

Reporting Period: September 1 2011 to August 31 2014

Overview: This project focused on development of novel derivative free optimization methods that rely on recent techniques and models from statistical learning. The main idea of these methods is to build local models of the objective function from randomly sampled data points. This approach has many benefits, in that it allows us to construct fairly accurate models with relatively small number of samples. The key difference with the deterministic sampling approaches is that these accurate models are constructed with some high probability, but not always. Moreover, it is not known, when these models are accurate. Only the probability of an accurate model occurring is known. Under these conditions, novel convergence theory needed to be developed, which has been the focus of our research. Below we list specific contributions.

Sparse Hessian models: We have developed theory and implementation for recovering sparse Hessian information in derivative free optimization. We showed that quadratic interpolation models computed by partial l1-minimization recover the Hessian sparsity of the function being modeled. Given a considerable level of sparsity in the unknown Hessian of the function, such models often achieve the accuracy of second order Taylor ones with very few random sample points. In particular, to construct an accurate second order model of an n-dimensional smooth function, in general , $O(n^2)$ local sample points are required. We have shown that if the Hessian of the function contains *s* nonzeros, then only $O((n+s)*\log(n)^4)$ samples are needed. Our results rely on compresses sensing theory and analysis of structured random matrices.
The results from this work have been the topic of an invited semiplenary at International Symposium for Mathematical Programming in 2012 and also the resulting paper was awarded the Informs Optimization Society Best Student Paper prize.

Convergence of trust region methods based on random models.
Due to the randomness in the sampling process the accuracy of the

models is also random, hence it was necessary to investigate convergence properties of trust region methods based on random models.

Traditional analysis of model based derivative free optimization methods relies on the worst-case behavior of the algorithmic steps and the models involved. There are conditions that the models and the iterates have to satisfy to guarantee convergence. Such requirements are difficult or costly to satisfy in practice and are often ignored in practical implementations. We developed a probabilistic view point for such algorithms, showing that convergence still holds even if some properties fail with some small enough probability.

In fact, using martingale theory we show that the probability of accurate models needs to be simply at least one half. This approach is first of its kind and has already generated follow-up research from other people in the field. In particular there are extensions to stochastic setting and convergence rates.

We also developed an implementation for noisy derivative free optimization for various conditions on the noise. We tested regularized models (Ridge regression and SVM models) in the derivative free setting. We extensively tested our software on a complex problem of protein alignment. The ridge regression models did not produce a noticeable improvement over the regular regression models. SVM models have shown improvements in some cases. The SVM work was performed with a visiting PhD student from a university in Brazil. He is continuing to work on this topic for his thesis.

Convergence of trust region methods for stochastic functions.
The ultimate goal of this project is development and analysis of trust-region model-based algorithm for solving black-box stochastic optimization problems. We proposed and analyzed a trust region framework, which utilizes random models of f(x). It also relies on (random, noisy) estimates of the function values at the current iterate to gauge the progress that is being made. The convergence analysis then relies on requirements that these models and these estimates are sufficiently accurate with sufficiently high probability. Beyond these conditions, no assumptions are made about how these models and estimates are generated. In the case when the estimates are accurate with probability one, our results recover the convergence

results for deterministic functions based on random sample sets, as described in in the previous section. Our method applies to both black-box and gradient based optimization and to both biased and unbiased noise. Our computational results show great advantage of using our framework over existing methods under different noise models.

We also developed a novel general approach for generating the sufficiently accurate random models by randomly sampling the objective function and constructing regression models (or other statistical learning models) based on these samples. Previously, models of stochastic functions relied on Monte-Carlo type (repeated) sampling of the function at deterministically selected points. Under these conditions it was possible to show that eventually an accurate model is obtained with increasingly high probability. With our work we are able to show that accurate models can be obtained by a much more general class of models and samples. Moreover, we show that is not necessary for the probability of an accurate model to increase. It can remain constant, as long as it is sufficiently high. Hence the number of sample points that are required does not need to grow as rapidly and previously used. Our computational results show the benefits of using random models and estimates. We have tested our approach on the (noisy) protein alignment problem, which has been one of our focus applications.

Archival publications resulting from this project:

Bandeira A., Scheinberg K., Vicente L.N., Computation of sparse low degree interpolating polynomials and their application to derivative-free optimization. Mathematical Programming 134(1) (2012), pp. 223-257.

Chen B., Chen R., Scheinberg K., Aligning ligand binding cavities by optimizing superposed volume. In Proceedings of IEEE BIBM 2012.

Bandeira A., Scheinberg K., Vicente L.N., Convergence of trust-region methods based on probabilistic models. SIAM Journal on Optimization (SIOPT), 24(3), pp. 1238-1264.

Chen, R., Menickelly, M., Scheinberg, K. Stochastic Optimization Using a Trust-Region Method and Random Models. Technical

Report. ISE Department, Lehigh University, 2015.

Cartis, C. and Scheinberg, K. Global convergence rate analysis of unconstrained optimization methods based on probabilistic models, Technical Report. ISE Department, Lehigh University, 2015.

Changes in research objectives, if any: none

Change in AFOSR program manager, if any: new PM, Fariba Fahroo